

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/793,408A

1636

DATE: 04/09/98
TIME: 14:45:15

INPUT SET: S24836.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

ENTERED

#23
JH
5-12-99

- 1
2
3 (1) General Information:
4
5 (i) APPLICANT: Choo, Yen
6 Klug, Aaron
7 Sanchez Garcia, Isidro
8
9 (ii) TITLE OF INVENTION: Improvements in or Relating to
10 Binding Proteins for Recognition of DNA
11
12 (iii) NUMBER OF SEQUENCES: 125
13
14 (iv) CORRESPONDENCE ADDRESS:
15 (A) ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
16 (B) STREET: 1100 New York Avenue, N.W.
17 (C) CITY: Washington
18 (D) STATE: D.C.
19 (E) COUNTRY: USA
20 (F) ZIP: 20005-3918
21
22 (v) COMPUTER READABLE FORM:
23 (A) MEDIUM TYPE: Diskette
24 (B) COMPUTER: IBM PC compatible
25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
26 (D) SOFTWARE: Word Perfect
27
28 (vi) CURRENT APPLICATION DATA:
29 (A) APPLICATION NUMBER: US 08/793,408
30 (B) FILING DATE: 02-JUN-1997
31 (C) CLASSIFICATION:
32
33 (vii) PRIOR APPLICATION DATA:
34 (A) APPLICATION NUMBER: PCT/GB95/01949
35 (B) FILING DATE: 17-AUG-1995
36
37 (viii) PRIOR APPLICATION DATA:
38 (A) APPLICATION NUMBER: GB 9514698.1
39 (B) FILING DATE: 18-JUL-1995
40
41 (ix) PRIOR APPLICATION DATA:
42 (A) APPLICATION NUMBER: GB 9422534.9
43 (B) FILING DATE: 08-NOV-1994
44
45 (x) PRIOR APPLICATION DATA:
46 (A) APPLICATION NUMBER: GB 9416880.4

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/793,408ADATE: 04/09/98
TIME: 14:45:16

INPUT SET: S24836.raw

47 (B) FILING DATE: 20-AUG-1994
48
49 (2) INFORMATION FOR SEQ ID NO: 1:
50
51 (i) SEQUENCE CHARACTERISTICS:
52 (A) LENGTH: 60 base pairs
53 (B) TYPE: nucleic acid
54 (C) STRANDEDNESS: single
55 (D) TOPOLOGY: linear
56
57 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
58
59 CTCCTGCAGT TGGACCTGTG CCATGGCCGG CTGGGCCGCA TAGAATGGAA 50
60
61 CAACTAAAGC 60
62
63 (2) INFORMATION FOR SEQ ID NO: 2:
64
65 (i) SEQUENCE CHARACTERISTICS:
66 (A) LENGTH: 92 amino acids
67 (B) TYPE: amino acid
68 (C) STRANDEDNESS:
69 (D) TOPOLOGY: unknown
70
71 (ii) MOLECULE TYPE: protein
72
73 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
74
75 Met Ala Glu Glu Arg Pro Tyr Ala Cys Pro
76 5 10
77
78 Val Glu Ser Cys Asp Arg Arg Phe Ser Arg
79 15 20
80
81 Ser Asp Glu Leu Thr Arg His Ile Arg Ile
82 25 30
83
84 His Thr Gly Gln Lys Pro Phe Gln Cys Arg
85 35 40
86
87 Ile Cys Met Arg Asn Phe Ser Xaa Xaa Xaa
88 45 50
89
90 Xaa Leu Xaa Xaa His Xaa Arg Thr His Thr
91 55 60
92
93 Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys
94 65 70
95
96 Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg
97 75 80
98
99 Lys Arg His Thr Lys Ile His Leu Arg Gln

PAGE: 3

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/793,408A

DATE: 04/09/98
TIME: 14:45:17

INPUT SET: S24836.raw

```
100                                     85                                     90
101
102      Lys Asp
103
104      (2) INFORMATION FOR SEQ ID NO: 3:
105
106          (i) SEQUENCE CHARACTERISTICS:
107              (A) LENGTH: 26 base pairs
108              (B) TYPE: nucleic acid
109              (C) STRANDEDNESS: single
110              (D) TOPOLOGY: linear
111
112          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
113
114      TATGACTTGG ATGGGAGACC GCCTGG                                     26
115
116      (2) INFORMATION FOR SEQ ID NO: 4:
117
118          (i) SEQUENCE CHARACTERISTICS:
119              (A) LENGTH: 28 base pairs
120              (B) TYPE: nucleic acid
121              (C) STRANDEDNESS: single
122              (D) TOPOLOGY: linear
123
124          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
125
126      AATTCCAGGC GGTCTCCCAT CCAAGTCA                                     28
127
128      (2) INFORMATION FOR SEQ ID NO: 5:
129
130          (i) SEQUENCE CHARACTERISTICS:
131              (A) LENGTH: 21 base pairs
132              (B) TYPE: nucleic acid
133              (C) STRANDEDNESS: single
134              (D) TOPOLOGY: linear
135
136          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
137
138      TATATAGCGT GGGCGTATAT A                                         21
139
140      (2) INFORMATION FOR SEQ ID NO: 6:
141
142          (i) SEQUENCE CHARACTERISTICS:
143              (A) LENGTH: 24 base pairs
144              (B) TYPE: nucleic acid
145              (C) STRANDEDNESS: single
146              (D) TOPOLOGY: linear
147
148          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
149
150      GCGTATATAC GCCCAGCTA TATA                                       24
151
152      (2) INFORMATION FOR SEQ ID NO: 7:
```

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/793,408ADATE: 04/09/98
TIME: 14:45:18

INPUT SET: S24836.raw

153
154 (1) SEQUENCE CHARACTERISTICS:
155 (A) LENGTH: 21 base pairs
156 (B) TYPE: nucleic acid
157 (C) STRANDEDNESS: single
158 (D) TOPOLOGY: linear
159
160 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
161
162 TATATAGCGN NNGCGTATAT A 21
163
164 (2) INFORMATION FOR SEQ ID NO: 8:
165
166 (1) SEQUENCE CHARACTERISTICS:
167 (A) LENGTH: 24 base pairs
168 (B) TYPE: nucleic acid
169 (C) STRANDEDNESS: single
170 (D) TOPOLOGY: linear
171
172 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
173
174 GCGTATATAC GCNNNCGCTA TATA 24
175
176 (2) INFORMATION FOR SEQ ID NO: 9:
177
178 (1) SEQUENCE CHARACTERISTICS:
179 (A) LENGTH: 33 base pairs
180 (B) TYPE: nucleic acid
181 (C) STRANDEDNESS: single
182 (D) TOPOLOGY: linear
183
184 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
185
186 TTCCATGGAG ACGCAGAAGC CCTTCAGCGG CCA 33
187
188 (2) INFORMATION FOR SEQ ID NO: 10:
189
190 (1) SEQUENCE CHARACTERISTICS:
191 (A) LENGTH: 33 base pairs
192 (B) TYPE: nucleic acid
193 (C) STRANDEDNESS: single
194 (D) TOPOLOGY: linear
195
196 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
197
198 TTCCATGGAG ACGCAGGTGA GTTCCTCAG CCA 33
199
200 (2) INFORMATION FOR SEQ ID NO: 11:
201
202 (1) SEQUENCE CHARACTERISTICS:
203 (A) LENGTH: 33 base pairs
204 (B) TYPE: nucleic acid
205 (C) STRANDEDNESS: single

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/793,408ADATE: 04/09/98
TIME: 14:45:19

INPUT SET: S24836.raw

206 (D) TOPOLOGY: linear
207
208 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
209
210 CCCCTTTCTC TTCCAGAAGC CCTTCAGCGG CCA 33
211
212 (2) INFORMATION FOR SEQ ID NO: 12:
213
214 (i) SEQUENCE CHARACTERISTICS:
215 (A) LENGTH: 33 amino acids
216 (B) TYPE: amino acid
217 (C) STRANDEDNESS:
218 (D) TOPOLOGY: unknown
219
220 (ii) MOLECULE TYPE: peptide
221
222 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
223
224 Met Ala Glu Glu Lys Pro Phe Gln Cys Arg
225 5 10
226
227 Ile Cys Met Arg Asn Phe Ser Asp Arg Ser
228 15 20
229
230 Ser Leu Thr Arg His Thr Arg His Thr Gly
231 25 30
232
233 Glu Lys Pro
234
235 (2) INFORMATION FOR SEQ ID NO: 13:
236
237 (i) SEQUENCE CHARACTERISTICS:
238 (A) LENGTH: 33 amino acids
239 (B) TYPE: amino acid
240 (C) STRANDEDNESS:
241 (D) TOPOLOGY: unknown
242
243 (ii) MOLECULE TYPE: peptide
244
245 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
246
247 Met Ala Glu Glu Lys Pro Phe Gln Cys Arg
248 5 10
249
250 Ile Cys Met Arg Asn Phe Ser Glu Arg Gly
251 15 20
252
253 Thr Leu Ala Arg His Glu Lys His Thr Gly
254 25 30
255
256 Glu Lys Pro
257
258 (2) INFORMATION FOR SEQ ID NO: 14:

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/08/793,408A

DATE: 04/09/98
TIME: 14:45:20

INPUT SET: S24836.raw

Line

Error

Original Text